

---

**Nucleotide sequence and expression of the cloned gene of bacteriophage SP6 RNA polymerase**

---

Hirokazu Kotani, Yukuo Ishizaki, Nobutsugu Hiraoka and Akira Obayashi

---

Central Research Laboratories, Takara Shuzo Co. Ltd, Seta 3-4-1, Otsu, Shiga 520-21, Japan

---

Received January 21, 1987; Accepted February 27, 1987

Accession no. Y00105

---

**Abstract**

The coding region of the gene for bacteriophage SP6 RNA polymerase was cloned into pBR322, and its entire nucleotide sequence was deduced. The predicted amino acid sequence for the polymerase consists of 874 amino acid residues with a total molecular weight of 98,561 daltons. Comparison of the amino acid sequence with that of T7 RNA polymerase reveals that regions with partial homology are present along the sequence. The coding region of SP6 RNA polymerase was inserted into an E.coli expression vector. The polymerase gene was efficiently expressed in E.coli cells, and the enzymatic properties of the expressed polymerase were very similar to those of the enzyme synthesized in SP6 phage-infected Salmonella typhimurium cells.

**Introduction**

Bacteriophage SP6 is a female specific DNA phage of Salmonella typhimurium (1). It seems that morphology and development of phage SP6 are similar to those of E.coli phage T7. The genomes of these two phages are in a linear, double stranded form and are comparable in size (43.5 Kb and 40Kb, respectively), but no extensive DNA sequence homology has been detected by the hybridization technique (2). Soon after infection of phage SP6, the phage development in cells becomes resistant to rifampicin due to the synthesis of phage-induced RNA polymerase (2). The SP6 RNA polymerase had been purified to apparent homogeneity. The polymerase consists of a single polypeptide with a molecular weight of around 100,000, as the analogous enzymes which are induced by other T7-like phages (2-5). The nucleotide sequence of T7 DNA is completely known, and the coding sequence for T7 RNA polymerase as well as genetic signals surrounding it have been assigned (6-8).

Recently, SP6 RNA polymerase was found to possess a very stringent promoter specificity which makes it distinct from those of analogous enzymes. Since then this polymerase is widely used for purposes such as preparation of precursor RNA and highly specific hybridization probes (9,10). Nevertheless, little is known about the genetical and structural properties of SP6 RNA polymerase. Purification of the polymerase from infected cells is difficult, because the enzyme is synthesized only for a few minutes after infection and does not accumulate to high levels.

To elucidate the organization, structure and expression of the SP6 RNA polymerase gene, we now cloned this gene and determined the coding sequence. In addition, expression experiments of this gene in E.coli cells were carried out.

### **MATERIALS AND METHODS**

#### **Bacterial strains, bacteriophage, and plasmid vectors**

E.coli HB101 (11) was used to harbor recombinant plasmids, E.coli JM109 for expression of the cloned SP6 RNA polymerase and M13 dideoxy sequencing, and S.typhimurium LT2 for the proliferation of phage SP6. Phage SP6 was the DNA donor for cloning. pBR322 (12), pUC18 (13) and pSP65 (10) were used for cloning, expression vectors and as template of enzyme assay, respectively. Phage M13 strains used for sequencing were mp18 and mp19 (13). S.typhimurium LT2 and phage SP6 were gratefully gifted from Y.Okada (Tokyo Univ.)

#### **Preparation of phage SP6 DNA**

S.typhimurium was grown at 40°C in L-broth supplemented with 6g/liter of Na<sub>2</sub>HPO<sub>4</sub> and 3g/liter of KH<sub>2</sub>PO<sub>4</sub> (modified L-broth). When the cell density reached  $2.4 \times 10^6$  cells/ml, SP6 phage was added at a multiplicity of 0.05 pfu/cell. Incubation was done at 40°C until the cell lysed. To the lysate was added chloroform (8ml/liter of culture) and the lysate was shaken for an additional 15 min. Phages were then concentrated according to the method of Yamamoto et al (14), and purified by pelleting through glycerol step-gradient ultracentrifugation. SP6 DNA was prepared from purified phage particles by phenol extraction.

#### **DNA manipulations**

Purification of plasmids, agarose gel electrophoresis, Bal 31

digestion, isolation, restriction, ligation of DNA fragments, transformation of cells, and DNA sequencing were mainly performed according to the procedures described by Maniatis *et al* (15) and Messing (16).

#### Expression of the cloned SP6 RNA polymerase gene in E.coli

*E.coli* JM109 containing the recombinant plasmid pSP6-2 was grown at 37°C in L-broth. When the cell density reached  $1.6 \times 10^8$  cells/ml, IPTG (isopropyl- $\beta$ -D-thiogalactopyranoside) was added to a final concentration of 0.2mM and cultivation at 37°C was continued for 3 hr. The cells were collected by centrifugation, washed with saline, and stocked at -70°C until the enzyme assay was conducted.

#### Preparation of extracts

The frozen cells were suspended in 5 volumes of buffer containing 50 mM Tris/Cl (pH 8) and 10 mM  $\beta$ -mercaptoethanol, and disrupted by ultrasonication. Cell debris was removed by centrifugation for 20 min at 8,000 rpm and 4°C, and the supernatant was used for assay of SP6 RNA polymerase.

#### Enzyme assay

The activity of SP6 RNA polymerase was assayed in a reaction mixture (100  $\mu$ l) containing 40 mM Tris/Cl (pH7.5), 6 mM  $MgCl_2$ , 4 mM spermidine, 10 mM DTT, 0.4 mM each of ATP, UTP, and CTP, 0.4 mM [ $^3H$ ]-GTP, 1  $\mu$ g template DNA and 1  $\mu$ l of the enzyme solution. Reaction was initiated by the addition of enzyme at 37°C. Incubation was usually done for 5 min and terminated by adding 1 ml of cold 5 % trichloroacetic acid (TCA). The precipitates were collected on Whatman GF/Cs filters and cold TCA-insoluble radioactivities were determined (17). One unit of the SP6 RNA polymerase activity was defined as the activity incorporating 1 nmol [ $^3H$ ]-GMP in one hour.

#### Other procedures

SDS-polyacrylamide gel electrophoresis (PAGE) was carried out by the procedure of Cox *et al* (18), and protein bands were visualized by silver staining. Dot matrix analysis was performed with the programs of DNASIS (Hitachi Software Engineering Co.,Ltd).

#### Other enzymes and chemicals

Restriction endonuclease, Ligation kit, M13 sequence kit and

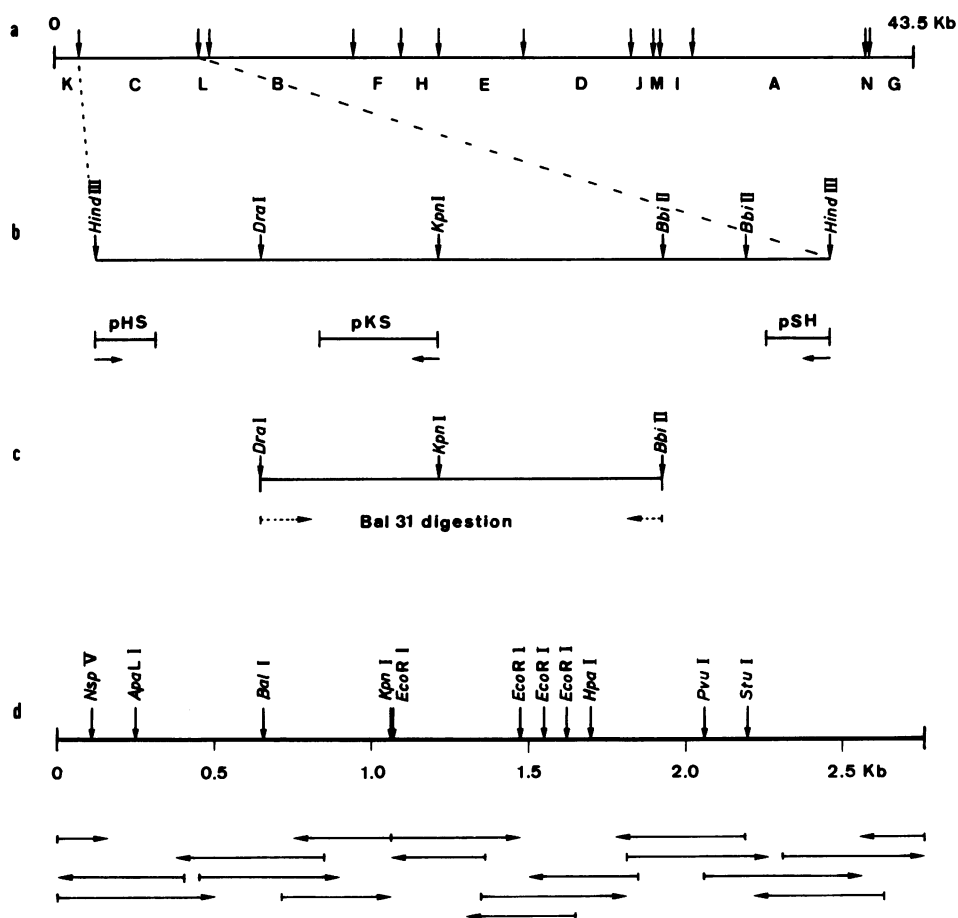


Fig. 1 Cloning and Sequencing strategy of the gene for SP6 RNA polymerase. (a) The *Hind* III restriction map of SP6 DNA. The arrows indicate *Hind* III sites and resulting fragments are named A to N according to the fragment sizes. (b) The restriction map of the 6.2 Kb *Hind* III C fragment. Under the map, the locations of the three clones (pHS, pKS and pSH) were indicated, and regions sequenced were shown by horizontal arrows. (c) The 3.4 Kb *Dra* I-*Bbi* II fragment used for cloning of the entire polymerase gene. The dotted arrows indicate the direction of Bal 31 digestion. (d) The restriction map and sequencing strategy of the inserted DNA segment in plasmid pSP6-1. Vertical arrows indicate restriction sites of enzymes, and horizontal arrows present the direction and regions sequenced.

sequence primers were supplied by our laboratories. [ $^3\text{H}$ ]-GTP (8.6 mCi/mmol) and  $\alpha$ -[ $^{32}\text{P}$ ]-dCTP (400mCi/mmol) were purchased from Amersham (England).

## **Results and Discussion**

### **Identification of the locus of SP6 RNA polymerase**

The Hind III restriction map of the bacteriophage SP6 genome is shown in Fig.1a (19). The early genetic region of SP6 development is presumably located on the left side of the SP6 DNA map, because E.coli RNA polymerase is known to preferentially transcribe in the left region of the SP6 genome (19). Thus, we assumed that the SP6 RNA polymerase gene is located in the Hind III-C fragment covering most of the early genetic region (see Fig.1a).

Based on this assumption, we first attempted to clone the Hind III-C fragment into pBR322, but no clone was isolated even under the conditions that Hind III fragments of similar size could be efficiently cloned. The result suggests the possibility that as in the case of phage T7, lethal sequences are located in the flanking regions of the polymerase gene. Then we used the following strategy to find the polymerase gene location within the Hind III-C fragment. The Hind III-C fragment was digested with Sau3A I plus Kpn I, and the resulting small fragments were cloned into M13 mp18. Among the colonies isolated, the three clones, pHS, pKS and pSH, of which the locations are shown in Fig.1b were selected, and nucleotide sequences of the portions indicated by arrows in Fig.1b were determined. Then the possible amino acid sequences were deduced, and homology to the T7 RNA polymerase sequence was searched by using the DNASIS programs. As a consequence, we found that the clone pKS carried a homologous sequence: The sequence of 61 amino acid residues deduced from the pKS nucleotide sequence showed 45% homology with the region of 308th-366th amino acid residues of T7 RNA polymerase (total amino acid residues=883) (8). Thus, we concluded that the gene of SP6 RNA polymerase resided on the 2.7 Kb segment which included the Kpn I site in the Hind III-C fragment.

Fig. 2 Entire nucleotide sequence of the cloned DNA segment and predicted amino acid sequence for SP6 RNA polymerase. Both the nucleotide and amino acid sequences were numbered from the initiation codon assigned. The SD-like sequence is underlined.

---

### Cloning of the SP6 RNA polymerase gene

According to the above results, the 3.4 Kb Dra I-Bbi II fragment was prepared from the Hind III-C fragment consulting the restriction map (Fig.1b). However, cloning of the Dra I-Bbi II fragment into the Cla I-EcoR V site of pBR322 was again unsuccessful, implying that lethal regions are present in the close vicinity of the polymerase gene. In the case of T7 phage, a weak E.coli RNA polymerase promoter immediately 5' proximal to the polymerase gene and two T7 RNA polymerase promoters in the 3' proximal region are shown to be lethal to E.coli cells (20,21). Thus, the Dra I-Bbi II fragment was partially digested with Bal 31 exonuclease to remove the presumed lethal regions as shown in Fig.1c, and was cloned into the EcoR V site of pBR322. About 50 ( $\text{Amp}^r$ ,  $\text{Tet}^s$ ) transformants were selected, from which plasmid DNAs were prepared, and the sizes of inserts were analyzed. One clone that apparently carried the longest DNA insert was isolated and named pSP6-1. This clone was submitted to restriction analysis, and its entire nucleotide sequence was determined according to the strategy indicated in Fig.1d.

### Nucleotide sequence of SP6 RNA polymerase

The nucleotide sequence of the cloned DNA segment in pSP6-1 is shown in Fig.2. The sequence consisted of 2,759 base pairs, of which the longest open reading frame was found to occupy 2,622 base pairs. This frame starts at the 32nd position and the first ATG sequence appears at the 41st position. We assigned this ATG sequence to the translational initiation codon, as it is accompanied by an SD-like sequence (22) AGGA in the appropriate position (see Fig.2). In Fig.2, the sequence was numbered from this ATG codon, and the predicted amino acid sequence was indicated above the nucleotide sequence. The peptide consists of 874 amino acid residues including f-methionine. This gives a molecular weight of 98,561 daltons which is well in agreement with the value of the polymerase from infected cells (100,000), as estimated by the SDS-PAGE analysis.

According to the above assignments, the polymerase gene is localized in the 3.3-6.0 Kb region from the left most end of the SP6 genome map. Since the T7 RNA polymerase gene has been mapped at the nucleotide positions 3,180-5,828 (6), the result reveals

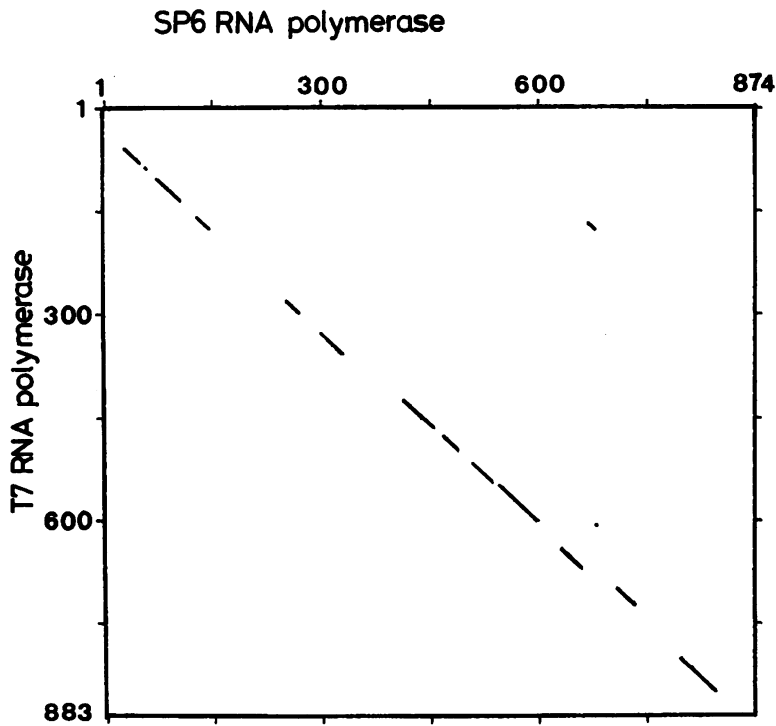


Fig. 3 Dot matrix comparison of SP6 and T7 RNA polymerases. Horizontal and vertical axes represent SP6 and T7 RNA polymerase amino acid residues, respectively.

Table 1.

	RNA polymerase activity (units/ $\mu$ l)				
	Templete	pSP65		pBR322	
	Rifampicin	-	+	-	+
JM 109		0.6	0.1	0.2	0
JM 109/pSP6-2 (IPTG -)		51.9	48.9	0.8	0
JM 109/pSP6-2 (IPTG +)		157.6	141.1	0.7	0.4
SP6 RNA polymerase		8.2	5.8	0	0
<u>E.coli</u> RNA polymerase		7.2	0	4.4	0

Expression of the cloned SP6 RNA polymerase gene in E.coli  
Enzyme assey was carried by using 1  $\mu$ l of crude extracts. As  
control, aliquots of purified SP6 and E.coli RNA polymerases were  
used.



that these two phages exhibit a marked similarity in gene organization.

#### Dot matrix analysis of SP6 and T7 RNA polymerase

The result of the dot matrix analysis between SP6 and T7 RNA polymerases is shown in Fig.3, in which the plot was done for every 30th amino acid position under the constraint that over 33 % homology was detected in a series of 30 amino acid sequences. It is apparent that partially homologous regions are present along the sequences. However, homology in each region is relatively low, and the regions with partial homology were flanked by non homologous regions which included some gaps. We interpret the result that the conserved homologous regions in both SP6 and T7 polymerases correspond to such functional domains of the enzymes as those involved in promoter recognition, substrate binding and polymerization. It should be mentioned that no significant homology was detected when comparison was made at the nucleotide sequence level (data not shown).

#### Expression of SP6 RNA polymerase in E.coli

To examine whether the cloned DNA segment actually encodes the functional SP6 RNA polymerase, we performed expression experiments. The cloned DNA segment does not contain the promoter-like sequence as shown in Fig.2, so that we placed the segment under the control of the lac promoter. The BamH I-Hind III fragment of pSP6-1 containing short pBR322 moieties on both sides of the cloned DNA segment was inserted downstream of the lac promoter of pUC 18. The ligated products were introduced into E.coli JM109 cells, and a colony carrying the recombinant plasmid (pSP6-2) was isolated. E.coli JM109 cells harboring pSP6-2 were cultured in the presence or absence of the inducer IPTG, and the cell lysates were prepared by according to the method section. When induced with IPTG, a high enzyme activity was detected as shown in Table 1. This activity was not affected by the addition of rifampicin. Only pSP65 DNA became the template, and no activity was detected with pBR322. This level of activity was about 100-folded higher than the enzyme level in S.typhimurium cells infected with phage SP6 (2). Without induction, the activity was about one third of that induced, and no activity was detected in the absence of the pSP6-2 plasmid.

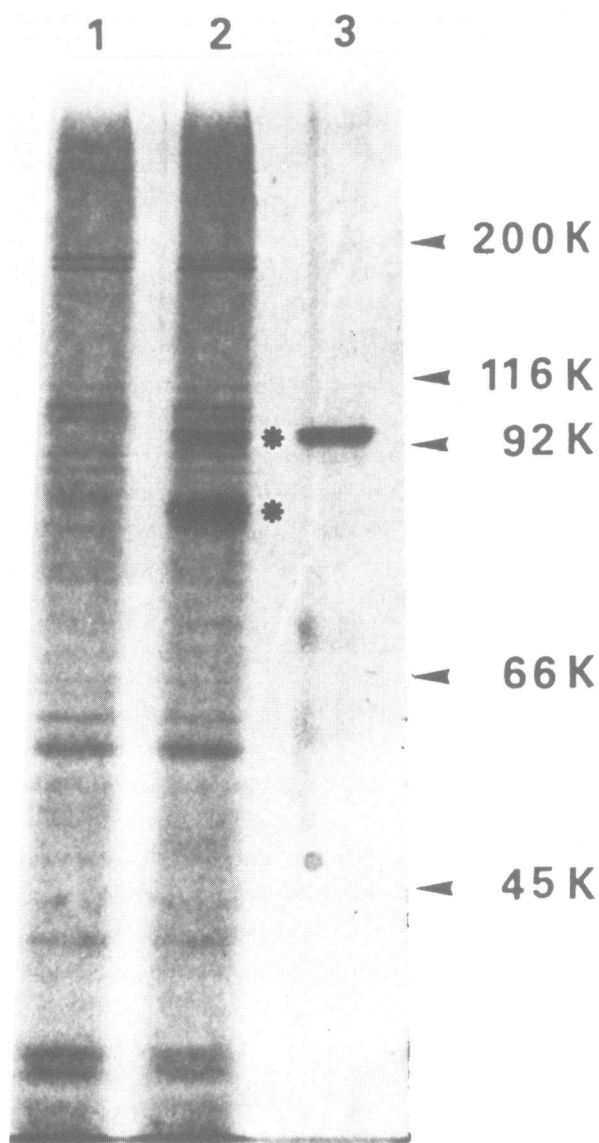


Fig. 4 SDS-polyacrylamide gel electrophoresis of crude extracts from cells harboring pSP6-2. Lane 1: non-induced extract, Lane 2: IPTG induced extract, Lane 3: authentic SP6 RNA polymerase as marker. The asterisks indicate the newly appeared bands by IPTG induction. The positions of size markers in Kilo-daltons are given by the side of columns.

The SDS-PAGE analysis of cell lysates with or without IPTG induction is shown in Fig.4. When induced by IPTG, two new major bands appeared at the positions of 100,000 and 83,000 daltons. The position of the former band exactly coincides with that of the authentic SP6 RNA polymerase. The later band is not identified yet, but it is assumed to be a proteolytically cleaved product of SP6 RNA polymerase during preparation of cell lysates, since a similar product has been observed when T7 RNA polymerase was overproduced (8).

Based on these observations, we concluded that the cloned DNA segment actually encodes the functional SP6 RNA polymerase and is efficiently expressed in E.coli cells. The nucleotide sequence is now available, and it is possible to produce a large amount of the enzyme in the E.coli system. We assume that the SP6 RNA polymerase system reported here will be very useful in elucidating the comprehensive structure and function of RNA polymerase by applying gene engineering techniques including site directed mutagenesis. It may also be possible to correlate respective functional domains of the polymerase to the conserved amino acid sequences.

Acknowledgement We thank Drs. H. Iba and Y. Okada of University of Tokyo and M. J. Chamberlin of University of California for gratefully supplying Salmonella typhimurium LT2 and phage SP6, and Dr. M. Takanami of Kyoto university for discussion and for reading of this manuscript.

#### References

1. Zinder, N. D. (1961), Science 133, 2069.
2. Butler, E. T., Chamberlin, M. J. (1982), J. Biol. Chem. 257, 5772-5778.
3. Chamberlin, M. J., McGrath, J., Waskell, L. (1970), Nature 228, 228-231.
4. Chakraborty, P. R., Sarkar, P., Hung, H. H. (1973), J. Biol. Chem. 248, 6637-6646.
5. Towle, H. C., Tolly, J. F., Boezi, J. A. (1975), J. Biol. Chem. 250, 1723-1733.
6. Dunn, J. J., Studier, F. W. (1983), J. Biol. Chem. 166, 477-535.
7. Stahl, S. T., Zinn, K. (1981), J. Mol. Biol. 148, 481-485.
8. Moffatt, B. A., Dunn, J. J., Studier, F. W. (1983), J. Mol. Biol. 173, 265-269.
9. Krieg, P. A., Melton, D. A. (1984), Nucl. Acids Res. 12, 7057-7070.

10. Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K., Green, M. R. (1984), *Nucl. Acids Res.* **12**, 7035-7056.
11. Boyer, H. W., Roulland-Dussoix, D. (1969), *J. Mol. Biol.* **41**, 459-472.
12. Boliver, F., Rodriguez, R. L., Greene, D. J., Betlach, M. C., Heynecker, H. L., Boyer, H. W. (1977), *Gene* **2**, 95-113.
13. Yanisch-Perron, C., Vieira, J., Messing, J. (1985), *Gene* **33**, 103-119.
14. Yamamoto, K., Alberts, B. M., Bevinger, R., Lawhorne, L., Treiber, G. (1970), *Virology* **40**, 734-744.
15. Maniatis, T., Fritsch, E. F., Sambrook, J. (1982), J. Eds, *Molecular Cloning*, Cold Spring Harbor Laboratories, Cold Spring Harbor, New York.
16. Messing, J. (1983), In: *Methods in Enzymol.* **101**, 20-78, Academic Press, New York.
17. Berg, D., Barrett, K., Chamberlin, M. J. (1971), In: *Methods in Enzymol.* **21**, 506-519, Academic Press, New York.
18. Cox, G. B., Downie, J. A., Fayle, D. R. H., Gibson, F., Radik, J. (1978), *J. Bacteriol.* **134**, 1141-1156.
19. Kassavetis, G. A., Butler, E. T., Roulland, D., Chamberlin, M. J. (1982), *J. Biol. Chem.* **257**, 5779-5788.
20. Davanloo, P., Rosenberg, A. H., Dunn, J. J., Studier, F. W. (1984), *Proc. Natl. Sci. USA* **81**, 2035-2039.
21. Tabor, S., Richardson, C. C. (1985), *Proc. Natl. Sci. USA* **82**, 1074-1078.
22. Shine, J., Dalgarno, L. (1974), *Nature* **254**, 34-38.